

DOCUMENT RESUME

ED 074 047

SP 006 273

AUTHOR Coleman, Peter
TITLE The Improvement of Aggregate Teaching Effectiveness
in a School Division. Occasional Paper No. 18.
INSTITUTION Manitoba Association of School Trustees, Winnipeg.
PUB DATE Dec 72
NOTE 37p.
EDRS PRICE MF-\$0.65 HC-\$3.29
DESCRIPTORS Behavioral Objectives; *Effective Teaching;
*Inservice Education; Teacher Behavior; *Teacher
Evaluation; Teacher Placement; *Teacher Selection;
*Teaching Quality

ABSTRACT

The aggregate level of teaching effectiveness in a school division can be improved over a 5-year period if attention is paid to the administrative decisions of teacher selection, assignment, development, and retention/release. A model has been constructed with four behavioral dimensions: warmth, indirectness, cognitive development, and enthusiasm. The empirical measurement of these dimensions allows predictive evaluation for teacher selection, formative evaluation for teacher development, and summative evaluation for retention or release decisions. The area of teacher assignment can be improved by consultation and student selection; however, more constructive methods are in the developmental stage. (The report contains four pages of references.) (BRB)

OCCASIONAL PAPER NO. 18

THE IMPROVEMENT OF AGGREGATE TEACHING
EFFECTIVENESS IN A SCHOOL DIVISIONABSTRACT

This proposal maintains that, over a period of about five years, the aggregate level of teaching effectiveness in a school division can be improved by careful attention to the administrative decisions of teacher selection, assignment, development, and retention/release. With the exception of assignment, good decision-making in these areas is dependent upon a sound teacher evaluation scheme.

Many such schemes have been proposed, and have been more or less ineffective, partly because no sound research basis existed for them. However, it is now possible to construct a model, the four behavioral dimensions of which have been validated by empirical studies. The dimensions of warmth, indirectness, cognitive development and enthusiasm have been operationally defined, have been measured, have been shown to relate positively to desirable attitudes and/or achievement levels in students, and can be developed in teachers. Additionally, at least two of them can reasonably be predicted. Thus these dimensions allow predictive evaluation for teacher selection, formative evaluation for teacher development, and summative evaluation for retention or release decisions.

The fourth area of administrative decision-making, teacher assignment, is less clear-cut. The match between teachers and assignments, based on training, experience, and preference can probably be improved by consultation. The match between teachers and students based on congruent expectations of role behavior (the students' mutual expectations for teacher behavior matching the teacher's expectations for teacher behavior, and vice versa) can probably be improved by arranging for student selection of preferred teacher(s), or by assessing expectations by a survey instrument. However, the gains here are hypothetical. As yet a sounder notion of constructive mismatch for growth along certain behavioral dimensions is also at the developmental stage.

Given appropriate data, and commitment to improving the quality of administrative decision-making in four vital areas, aggregate teaching effectiveness can be improved to the great advantage of the students in a school division.

INTRODUCTION

The improvement of teaching effectiveness of a school division, in the aggregate is, it will be maintained, a long term project requiring commitment over a period of about 5 years. This proposal involves four related activities each of which contributes to the increasing of the aggregate effectiveness of the teaching staff. The first activity, teacher selection, contributes to aggregate effectiveness by attempting to eliminate teacher candidates who would be ineffective in practice. The second activity, teacher assignment, attempts to improve aggregate effectiveness by improving the match between teachers and assignments. Teacher development, through self-help techniques and more formalized in-service processes, is an important contributor to staff effectiveness, at least potentially. The final element or activity is also a significant contributor to aggregate effectiveness, although not commonly considered as such. That is the release of teachers who have demonstrated to be relatively ineffective, and for whom the self-help and format in-service programs have not been useful. It is maintained that only such a comprehensive plan, persevered with over a period of years, can significantly affect the level of teacher effectiveness in a school division.

One important reservation should be entered here. All of the suggestions to be made here have been demonstrably effective in other contexts, and used for different purposes. But most of the activities, techniques, and instruments described here have never been used for the purposes suggested here. This does to some extent reduce the amount of

confidence one can have in the probable outcome. For example, the relationship between the verbal facility test scores of teachers and pupil achievement has been demonstrated several times. Yet several researchers have cautioned against using this demonstrated relationship, which is not necessarily causal, for administrative purposes. However, the probable cost-effectiveness of selecting teachers on this basis is extremely high compared to some other bases which have been used in the past. Thus the benefits seem to exceed the risks, in this particular case, and this particular technique will be proposed for administrative uses. Given the general reservation then, that many of the suggestions made here are based on relatively recent research findings, and are hence untried for the kinds of uses proposed, the suggestions made here seem well worth the consideration of practicing administrators, especially when it is considered that present teacher evaluation, and hence effectiveness, schemes have been very unproductive.

The arrangement of the paper is as follows: some general questions raised by the notion of effectiveness in teaching will be considered briefly, and then a general teacher evaluation model will be described. The four dimensions of the model and their application are then analyzed in some detail. The next section deals with the question of teacher assignment, which is somewhat different from the issues of selection, development, and release in that evaluation is not of fundamental importance. The final two sections suggest some of the problems of implementing such a proposal, and draw some conclusions and further implications from the material already presented.

Towards More Effective Teaching

Any attempt to improve the effectiveness of teaching must recognize the relative lack of success of many thousands of such attempts which have been made in the past. Caution and precision are hence essential, and any proposals must be relatively tentative and can only be justified by the importance and urgency of the task. In this proposal teaching is defined as "the exertion of the behavioral influence". (Gage, 1972: p. 43) Obviously behavioral influence can be positive or negative, that is it can change behavior in desired directions or non-desired directions, and it can also be more or less effective, that is it can make large scale or small scale changes. More effective teaching then is defined as the exertion of behavioral influence in desired directions and in larger measure than before.

The ways in which this influence is exerted on students are, as will be described in detail later, through various teacher behaviors in the classroom. Thus the notion that teaching can become more effective assumes that the behavior of teachers in classrooms can be changed. This is part of a general assumption about human behavior which is somewhat unpopular currently. In a recent article in a popular magazine, entitled "Human Beings Are Not Very Easy to Change After All: An Unjoyful Message and its implications for Social Programs," Etzioni suggests that social scientists "have begun to re-examine our core assumption that man can be taught almost anything and quite readily". (1972: p. 45) However, there

is a substantial body of research which demonstrates that teacher behavior can in fact be changed, at least in the context of formal teacher training programs. For example, Turner (1963) demonstrated that the improvement of teacher problem solving through training was reflected in the increases in achievement of pupils over periods of several years. More generally, a whole series of micro-teaching activities have demonstrated that teacher classroom behavior can be changed, and the changes resulting are fairly stable. (See Berliner, 1969)

A second major assumption being made here is that enough is known about the dimensions of effective teacher behavior to enable us to decide what represents more effective behavior. There is some real doubt about this.

The value of most teacher effectiveness studies in the past has been limited by their attempts to predict teaching success directly from assessments of the personal characteristics of teachers without considering any intervening variable, by their unreliable assessment of rather gross features of classroom behavior, and by their lack of attention to the varied contexts in which teachers work...the questions to which investigators were confidently seeking answers half a century ago can now be seen as answerable only on the basis of a great deal of research which has hardly begun. (Morrison & McIntyre, 1969: p. 41)

However, recent research seems to justify more optimism. The most recent review of the research on "teacher effectiveness" in the Encyclopedia of Educational Research suggests that

the research which is reviewed herein permits cautious optimism and indicates that the tools long needed for the analysis of the teaching-learning process are gradually being developed. This optimism is in contrast with the conclusions reached in past reviews. (Flanders & Simon, 1969: p. 1423)

The tentativeness of the conclusions arrived at so far in research on teacher effectiveness, and the lack of clarity about their applicability to specific school division situations, suggests the need for caution. However, the practical problem of evaluating teachers is so pressing that some uses of the research findings, however uncertain and tentative, seem essential. There is for the first time a reasonably well established set of research findings on teacher effectiveness, they do offer some possibility of use in practice, and the main purpose of the paper is to describe briefly these findings, and suggest how they might be used.

The dimensions of teacher behavior used here were developed originally by Gage, in his recent survey of research on teacher effectiveness. By surveying empirical research, he tentatively identified four dimensions of teacher behavior which seemed desirable. His procedure was to

Present a series of operational definitions of teacher behaviors that seem, more or less, to belong in the same dimension. These definitions will be drawn from various research procedures and measuring instruments. Then I shall cite some of the evidence on which it is possible to base the inference that these behaviors or characteristics are desirable. (Gage, 1972: p. 34)

The four dimensions he identified are "warmth", "indirectness", "cognitive organization", and "enthusiasm". As Gage points out, there is nothing very startling about these four variables. They do

not exhaust the possibilities and they are really only representative of the things that research on teaching can presently support. What is important about these dimensions "is their basis in empirical research". The ease with which others have told us such truths in the past is matched by their untrustworthiness." (p. 39) To summarize the reasons for selecting these dimensions in very brief form, first, they are based on empirical research; second, reliable instrumentation for measuring these teacher behaviors generally exists; third, the desirability of these teachers' behaviors has been demonstrated; and fourth, these behaviors can in fact be learned by teachers.

The approach has of course some major defects as well. It is far from comprehensive, and ignores a great many significant kinds of evaluations of school effectiveness. However, recent attempts at global evaluation systems, and particularly those based on student achievement, have all been to some extent unproductive. Tumin (1970) in an extensive examination of the issue of evaluating the effectiveness of education, is as pessimistic as he is thorough. He provides a series of questions which remain to be answered before a satisfactory model of evaluation can be constructed, (given in abbreviated form):

- "1. Whose goals should be taken as the goals of education in any system?
2. Assuming one knows somehow whose goals should be taken, how does one discover these goals?

3. How does one decide what outcomes shall be taken as proof of the accomplishment or failure to achieve those agreed upon goals?
4. How does one measure those outcomes?
5. How does one assess the contributions of various factors to the achievements so measured assuming one can measure?
6. How does one relate behavior and measures at the end of school careers to conduct and achievement in adult life?
7. How does one estimate the economic costs and the economic gains of various input and outcomes of the educational process?
8. How does one add up such analytically disparate outcomes as cognitive adequacy, emotional well-being, creativity, readiness for inter group life, ability to function in a democracy, readiness to change, and appreciation of the cultural traditions of the society?
9. Against what standards does one compare a school's "achievements?"

The validity of these observations is perhaps attested to by a number of recent failures to provide useful evaluations of the work of teachers.

For example, the limitations of using standardized tests of pupil achievement for making decisions about teachers have been emphasized by recent concerns with performance contracting. The difficulties of using standardized tests to provide data for definitive judgments about the effectiveness of performance contracting and the rewards due the contractors (Klein, 1971) are effectively the same as those involved in making judgments of teachers.

Another interesting recent failure to measure teacher competence in terms of student achievement was based on the teaching of technical

skills, which seem readily measurable outcomes, and used "performance tests". (See Klein and Alkin, 1972 for a discussion). Performance objectives were carefully defined, 10 hours of instruction were given by two groups of people, experienced teachers and non-teachers, and the outcome in terms of class achievement was measured. The results did not discriminate between the experienced teachers and the non-teachers. This approach presumably then could not in any way assist in discriminating degrees of teacher competence, since it cannot even distinguish between experienced and inexperienced teachers. (Popham, 1968)

The surprise or disappointment frequently expressed in this connection is itself remarkable, since a relatively early and extremely thorough study of the characteristics of teachers concluded that

Product measurements (estimates of the behavior or achievements of the pupils of teachers) have been widely acclaimed as desirable criterion data, but have been infrequently used in the study of teacher behavior. Actually, the seeming relevance and appropriateness of the measurement of pupil behaviors and their products as indicators of teacher performance may be more apparent than real, for the producers of (or contributors to) pupil behavior or pupil achievement are numerous, and it is difficult to designate and parcel out the contribution to a particular "product" made by a specified aspect of the producing situation, such as the teachers. (Ryans, 1960: p. 375)

This is of course virtually the identical conclusion arrived at after a great deal of discussion and debate with regard to the Coleman Report. Mood suggests that "the present rudimentary state of our quantitative models does not permit us to disentangle the effects of home, school, and peers on student achievement." (1970: p. 6) Specifically with

reference to teachers, Mood goes on to say "we can only make the not very useful observation that at the present moment we cannot make any sort of meaningful quantitative estimate of the effect of teachers on student achievement." (p. 7)

If it is granted that the global evaluation schemes had turned out to be less than satisfactory, then it seems reasonable to base evaluation of teachers on somewhat different measures. One way of interpreting the enormous quantity of research on teacher effectiveness since the last 30 or 40 years is that it too demonstrates the necessity for precision in evaluating teachers. "Where the earlier efforts made much use of global ratings, the present day work relies much more on reliable counts of specific behaviors". (Gage, 1972: p. 206) Hence what will be proposed here is a fairly specific set of dimensions, on the basis of which teachers can be evaluated directly. In due course as some of the difficulties with using standardized scores of student achievement are eliminated it may be possible to add that dimension to the teacher evaluation scheme proposed here for school divisions; at present, however, student achievement scores seem not to provide a realistic means of evaluating teachers in the context of school divisions.

Evaluating Teacher Effectiveness in a School Division

The current status of teacher evaluation programs in large school systems in the United States was reported recently by the Educational Research Service of the American Association of School Administrators (1972). Based on a survey of school systems in the United States enrolling 25,000 or more pupils, and on returns from 155 systems, the survey showed that the four major purposes of evaluations are "to stimulate improvement of teacher performance", "to decide on reappointment of probationary teachers", "to recommend probationary teachers for permanent status", and "to establish evidence where dismissal from service is an issue". To a question on the status of teacher evaluation procedures in 1971-1972, (155 answered). 110 systems noted that their program will remain unchanged, 35 noted that their program will be revised in 1971-72, 6 suggested that they did not have a program at present but would initiate one in the 1971-1972 school year, and 4 suggested that they had no plans for a teacher evaluation program. To a question regarding who was responsible for evaluating (108 answered), 77 of the responding systems noted that the principal was the sole evaluator responsible. In the overwhelming majority of school systems, classroom observations of teachers by principals or supervisors are the standard method of evaluating. In a majority of cases, a check list or rating form is used. On the basis of this survey, it seems reasonable to conclude that most large school

systems in the United States still make use of evaluation schemes the basis for which has been rather thoroughly discredited over a substantial period of time. Morrison and McIntyre summarize the case against rating scales thus:

Despite their popularity several objections can be raised against rating scales. One of their more serious limitations when used for assessing the classroom behavior of teachers is that an extensive amount of information about what has gone on has to be reduced to subjective and impressionistic endorsements on a few scales. Since they are heavily dependant upon the subjective impressions formed by the individual rather than their reliability from one occasion of rating to another by the same rater, or between two or more raters on the same occasion, is highly variable. Also, when the rater is presented with several supposedly distinct characteristics to assess he may in fact be unable to distinguish between them, leading to a tendency to rate an individual as 'high', 'average' or 'low' on most of them. Finally the information available to the rater can vary very much from one characteristic to another and from one individual to another. (1969: p.22)

This is not to say that ratings of teachers are necessarily and invariably inaccurate. However, unless the set of guidelines proposed by Ryans (1960: p. 75) or something similar is observed, those ratings will not be very reliable. The conditions are not normally observed in actual teacher rating systems, as the ERS Survey shows. The rating technique has never been acceptable to teachers, who have strongly resisted, via their professional associations, the suggestions that either their salary or their job security should be determined by ratings. Their success in achieving job tenure and pay scales unaffected by judgments of their competence arrived at through classroom visitations is clear evidence of the unspoken agreement by virtually all concerned that classroom visitation and

rating is neither a reliable nor a valid means of evaluating teachers.

The model proposed here satisfies the definition of educational evaluation adopted by the Center for the Study of Evaluation at UCLA which suggests that educational evaluation is

the process of determining the kinds of decisions that have to be made; selecting, collecting, and analyzing information needed in making these decisions; and then reporting this information to appropriate decision-makers. (Klein, 1971: p.9)

The three kinds of decisions which the model proposed here provides information for are: teacher selection, teacher development, and teacher release. Different kinds of information are required in each of these different decision areas, and thus the model provides for predictive, formative, and summative evaluation.

The model is summarized in the following chart, and subsequent sections of the paper elaborate on the four dimensions of teacher behavior used in the model.

THE EVALUATION OF TEACHER EFFECTIVENESS:

AN ADMINISTRATIVE MODEL

<u>Dimensions of Teacher Behavior</u>	<u>Types of Evaluation Yielded</u>	<u>Decision Area</u>
WARMTH	Predictive Formative	Selection Development
INDIRECTNESS	Formative Summative	Development Retention/Release
COGNITIVE DEV.	Predictive Summative	Selection Release
ENTHUSIASM	Formative Summative	Development Retention/Release

The Dimension of "Warmth"

The dimension which Gage labels "warmth" has been identified by three different instruments, the Minnesota Teacher Attitude Inventory, the California Scale, and the Teacher Characteristics Schedule, and these three instruments correlate fairly closely. Furthermore, all of the instruments identify attitudes and behaviors which correlate positively with favorable assessments of the teachers by both students and objective observers, and with the achievement of students. As Gage summarizes the situation,

A substantial body of evidence supports two conclusions:
a) teachers differ reliably from one another on a series of measuring instruments that seem to have a great deal in common. b) these reliable individual differences among teachers are fairly consistently related to various desirable things about teachers. (Gage, 1972: p. 35)

If the desirability of warmth is accepted, and it seems incontrovertible the various instruments described here can then be used predictively, that is as selection devices in hiring teachers, with reasonable probability that the teachers with favorable scores will be effective teachers, both in terms of students attitudes towards them, and the achievements of students.

In addition, there are further possibilities inherent in these research findings on the dimension of warmth. Since the M.T.A.I. has been shown to relate quite closely to favorable ratings of the teachers by their pupils (Yoss, 1967), and since student ratings of teachers have

been shown to be extremely useful in changing teacher behavior, this dimension can also provide formative evaluation.

Student ratings of teachers "have yielded useful evidence on teaching at levels as low as fourth grade and as high as medical school". (Gage, 1972: p. 172) In a long series of experiments in feedback from student teachers, a number of researchers have demonstrated that student feedback, particularly in the form of written ratings, is a reliable and valid way of helping teachers improve their teaching, at least as perceived by students. The improvement here does not generally reflect in gains in student achievement, but only in more positive attitudes in students. This technique, it should be emphasized, is really a private transaction between students and their teachers, and has no validity whatsoever for administrative uses, in for example evaluating teachers for retention or release. Nevertheless, it is a form of teacher development which has had good effect in a number of different contexts and could well be an important element in in-service training of teachers, provided the limitation on its use is clearly spelled out.

The Dimension of "Indirectness"

Gage bases this dimension on two different but relatively closely related research areas. The first is usually identified as "interaction analysis", and is associated with Flanders. One characteristic of teacher verbal behavior has been found in close association with both constructive student attitudes and favorable student achievement levels, in a number of studies by different researchers at different grade levels, and in different parts of the United States.

The percentage of teacher statements that make use of ideas and opinions previously expressed by pupils is directly related to average class scores on attitude scales of teacher effectiveness, liking the class, etc. as well as to average achievement scores adjusted by initial ability. (Flanders & Simon, 1959: p. 1426)

The variety of studies, and of ways in which the interaction analysis data has been used, as reviewed by Flanders and Simon, suggests that such data enables one to predict the relative success of teachers who fall into high or low categories in some types of verbal behavior, to train teachers to exhibit the apparently desirable verbal behavior, and to distinguish classrooms in which achievement and attitudes of students will be relatively poor. Thus this dimension has predictive, formative, and summative, and consequently this dimension would be used to develop in-service training programs for teacher development, and in evaluating teachers for retention or release.

The Dimension of "Cognitive Development"

This dimension of effective teacher behavior is certainly the least understood and least immediately useful way of measuring teacher effectiveness. Since it has not yet been satisfactorily defined operationally, it will be necessary here to suggest the use of a number of proxies, or indicators, in place of cognitive development directly measured by some reliable and valid observations or a test instrument. Unfortunately, these proxies are not "process variables" or even clearly related to such variables, and consequently do require the inference that the characteristics measured do affect teachers' classroom behavior. They have all been shown to correlate

positively with student achievement levels; what is in question is the degree of causal relationship. Proposed are first, verbal facility; second, academic achievement; and third, recency of academic training. Each of these proxies requires some description and specification.

The importance of the teacher's verbal ability was clearly demonstrated by the Coleman Report.

The most significant school service variable in explaining student achievement (measured by a vocabulary test) was a teacher characteristic, the teacher's verbal ability. As with the other findings of this nature that we have discussed care must be used in interpreting the meaning of such results ...However, if one views teacher's verbal ability as the proxy measure for a number of related skills and qualities, the Coleman Report finding can be interpreted in a meaningful fashion. If the measure of verbal ability is taken to represent the general intelligence level of the teacher, the finding can be construed to mean that an intellectually facile instructor is more adept at tasks such as finding means to motivate students, adapting materials to their ability levels, and communicating in ways which make the subject matter more understandable. This is an interpretation which is totally consistent with the observations and conventional wisdom of untold thousands who have themselves been teachers or who have supervised teachers. Guthrie, (1970: p. 37)

Although the relationship between teacher verbal ability and student achievement is not currently questioned, since the findings of the Coleman Report, the usefulness of this measure of teachers for administrative purposes is certainly debatable. Thus for instance Gage points out that this relationship is correlational, but not necessarily causal. Consequently, "we cannot proceed to improve student achievement by hiring teachers with greater verbal ability". (1972: p. 33)

Similarly, Mood points out that:

If we went about increasing the verbal ability of teachers, the increase that might result in student achievement would be far less than what would be calculated by using the equation that relates to achievement. The reason is that a specific increase in verbal ability would probably not be accompanied by a corresponding increase in all the other attributes that verbal ability is serving as a proxy for. (1970: p. 3)

However, other writers disagree. For example, Levin suggests that:

Recruiting and retaining teachers with higher verbal scores is five to ten times as effective per dollar of teacher expenditure in raising achievement scores of students as the strategy of obtaining teachers with more experience. (1970: p. 24)

The relationship between academic achievement of teachers and the achievement of students has been shown by many different studies at various times, and will not be further described. However, the finding that recency of academic training is more closely associated with student achievement than the overall level of training obtained is relatively new, and is an outcome of a re-analysis of Coleman Report data. Hanushek found that

Recent educational experiences - either undergraduate or graduate level - are important. Thus efforts to have teachers returned to school during summers seem justified in terms of effects on education. The cumulative effect (the master's degree or total units) is not as important as recent involvement. (1970: p. 92)

In summary, the dimension of "cognitive development" in teachers would be assessed by the development of three proxies, each of which has been shown to be associated with student achievement. It seems reasonable to infer that teachers with high relative scores on these

proxies would then be relatively more effective than other teachers. It is suggested that the administrative use made of these proxies would be both predictive and summative. That is, at the teacher selection phase teachers would be selected in part on the basis of verbal ability scores, academic achievement, and recency of academic training. In the evaluation for retention or release phase, similar use would be made of the proxy scores. Since these scores do not lend themselves very readily to the development of in-service training programs, they have no formative usefulness.

The second proxy, ability to explain, arises out of a series of experiments and studies by Gage and colleagues based on the notion that teaching behavior consists of a group of technical skills, amongst other things, and that one such skill was the ability to explain. Studies of this ability in teachers showed that the differential effectiveness of teachers as explainers was perceptible to trained observers using a video and audio record of the explanations given by the teacher in the classroom. More interestingly, from the point of view of teacher development, a manual developed by Miltz on "how to explain" which was based on the previous research, increased the ability of teacher trainees to explain.

The Dimension of "Enthusiasm"

The distinction between "enthusiastic" or "stimulating" teaching and "indifferent" teaching has been used in a substantial number of research studies. In a recent review of these studies Rosenshine (1970)

distinguished between high inference studies, which "require considerable inferring from what is seen or heard in the classroom to the labelling of the behavior" (p. 500) and low inference studies in which specific behaviors are carefully identified. The six high inference studies reviewed

provide strikingly consistent results. They suggest that one of the patterns of effective teaching behavior identified by Ryans (1960), namely Pattern Z, described as "stimulating, imaginative, surgent versus dull, routine teacher behavior" is significantly related to pupil achievement. (p. 506)

Reviewing the low inference studies which are particularly useful for teacher training programs because they do give some indication of the specific behaviors which teachers should exhibit, Rosenshine identified a number of behaviors as components of "enthusiasm", and hence desirable:

The teacher who scored high on the Energy factor appears to exhibit three types of related behavior. First, he is energetic, a rapid speaker, mobile and enthusiastic, but relaxed. Second, he asks varied questions, emphasizing questions of interpretation and opinion as well as factual questions. Third, he praises frequently. (p. 506)

Clearly, the high inference studies above relate more generally to the summative evaluation in the evaluation model here and the low inference studies relate to the formative evaluation stage. There seems little doubt however that the same general teacher behavior is being noted in both types of studies. Rosenshine's final summary of the studies on enthusiasm suggests that:

The results of high inference studies provide evidence that ratings given to teachers on such behaviors as "stimulating", "energetic", "mobile", "enthusiastic", and "animated" are related to measures of pupil achievement. The results of low inference studies suggest that the frequencies of such variables as movement, gesture, variation in voice, and eye contact are related to pupil achievement. (p. 510)

There is another research area which seems fairly closely related. Presumably enthusiastic and energetic teachers have differential impact on the attention level of the students. As Rosenshine points out, the result of the study described "may occur because animated behavior arouses the attending behavior of pupils", (p. 510) and in fact the attention level of pupils is itself an area of investigation with a long history. This has been reviewed and summarized by Jackson:

What, then, are the chief conclusions to be derived from systematic studies of classroom attention, starting from the early work of Morrison and extending to the most recent reports? First, although the amount of attention may vary considerably from class to class and even from minute to minute within a class, it would seem that most of the time most students are attending to the content of the lesson. Second, the amount of attention in the classroom is often less than meets the eye...Third, the amount of attention even when crudely estimated by an outside observer, seems to be significantly related to other educational variables, such as scores on achievement tests and estimates of teacher effectiveness. There is also the suggestion that the amount of attention may not be closely related to the students' intellectual ability. In sum, these conclusions provide ample justification for further study and speculation. (1968: p. 102)

Given the existence of the relationship between enthusiastic teaching and the attention of pupils which is suggested by the research, and which certainly seems consistent with common sense and the experience of practitioners, the use of observer estimates of the attention of the class to the teacher does not seem an unreasonable evaluation technique. Furthermore,

An estimate of pupil attention is commonly used by teachers to judge their personal effectiveness in the classroom. The possibility of massive inattention, signaling the loss of the

teacher's authority, is frequently reported as a dominant fear among beginning teachers. Second, students also worry at times about their inability to remain focused on the task at hand. Boredom is one of the chief complaints of students who are having difficulty with school. (Jackson, 1968: p. 102)

Thus there seems to be substantial justification for the use of observation of pupil attention as an element in formative evaluation as well. If the anxieties of teachers, particularly beginning teachers, and the boredom of students can be somewhat relieved by training teachers to be more enthusiastic or energetic, as the research suggests is possible, this could be a most useful dimension of teacher effectiveness.

Teacher Assignment: Improving the Match

The general issue of match or consistency between teachers and their work assignments can be considered in three different ways. First, the relatively simple question of the wishes of the teacher with regard to a teaching assignment; second, the more complex question of role expectations and role conflicts for teachers and students and third, the even more difficult, and relatively unexplored question of what is the most productive match (or mismatch) between students and teachers for promoting student growth along psychological, social, or cognitive dimensions can be considered.

A recent study of the teacher workforce in British Columbia (Wallin, 1971) showed that only 57.47% of the teachers responding (n=16,387) expressed satisfaction with their teaching assignment. The study built up a "quality of assignment" Index, of which the negative, or misassigned category, included a teacher who reported that he (she)

"(1) was dissatisfied with his present position,

- (2) found none or few formal courses in his pre-service education to be relevant to his current assignment,
- (3) indicated that his training assignment did not match, and
- (4) has not participated in any non-formal credit courses that pertain to his present assignment."

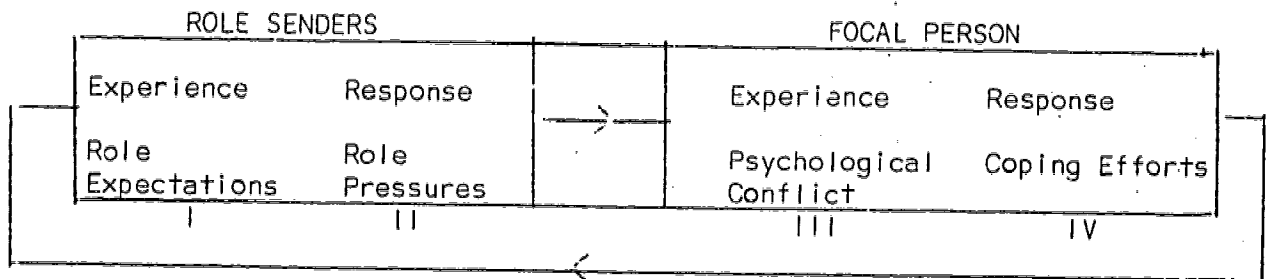
The positive or well-assigned categories included the teacher who reported that he

- "(1) wanted to have his present assignment next year, was satisfied with his assignment,
- (2) has had relevant formal courses in his pre-service education, or
- (3) has taken non-credit type courses and programs related to his present assignment."

The study found that perceived misassignment was relatively common, with 12.66% of teachers falling into this category, and that teachers who felt well-assigned included 26.46% of the workforce. Misassignment seemed to be particularly common amongst rather well-trained teachers.

This data on assignment has two different kinds of significance - first, it seems to indicate a concern amongst teachers regarding the misuse of talents, abilities, or training, and second, it probably also illustrates a resentment over lack of control of work, which is common amongst professionals working in organizations, including teachers. (Corwin, 1970: p. 46) Thus a reduction of this misassignment requires two kinds of solutions: a process of consultation, to increase the sense of control, and also a serious attempt by administrators to reduce mismatch between desired and actual assignments, and hence to alleviate concerns about the misuse of talents, abilities, or training.

Another version of matching can be based on the research on role behavior, particularly in organizational contexts. The theoretical model of factors in role conflict and ambiguity used here (see diagram) is that of Kahn et al, (1964).



The model is based on the notion of a role episode, "a complete cycle of role sending, response by the focal person, and the effects of that response on the role senders." (p. 26) In practice, conflict and ambiguity is experienced by the focal person through a series of such episodes. The notion of matching applied here suggests that appropriate teacher assignment might be able to minimize role conflict by ensuring a relatively high degree of congruence between student expectations and teacher expectations of appropriate role behavior. One caution is in order: although studies of expectations are common, studies of teacher-role performance are relatively uncommon, and it is quite possible that the continuity between role expectations and role performance will not be direct and straightforward, so that knowledge about expectations will not necessarily yield knowledge about classroom behavior. However, it seems possible that if there is a good match between student and teacher expectations, that these expectations might more readily be translated

into the expected behavior. For a thorough review of the sociological research in role theory, as it has been utilized in research on teachers and teaching, see Biddle (1969).

In order to achieve congruence in assignment, prior knowledge regarding the expectations of the teacher and the students is necessary. Although a good many test instruments for ascertaining role expectations of teachers and student exist, (see e.g. those used by Bogen, 1964, and Cheong and DeVault, (1966) it would probably be necessary to develop a simple test for the purpose.

Another technique, which might be used together with or instead of the pencil-and-paper test of expectations, is simply to offer students a choice of teachers and assume that they possess enough knowledge to choose the teacher whose behavior will conflict least with their expectations. The obvious and common-sense reaction amongst educators to this suggestion is that the students would all wish to be taught by the least demanding teacher. However, as is often the case, research does not support the glib generalizations of common sense. Both American and British studies have shown that there is a high degree of unanimity in what children regard as desirable qualities of teachers. Additionally, "children, especially younger children, described 'the good teacher' more in terms of his teaching than did teachers or student teachers; the latter group in particular emphasized the personal qualities of teachers". (Morrison & McIntyre, 1969: p. 109) The same British study being quoted above

found

Wide agreement among pupils about the importance of firmness, justice, the avoidance of corporal punishment, friendliness, knowledge and, most of all, participation in class activities. Amongst junior school pupils, encouragement to work hard was stressed, as was politeness by the girls; and amongst secondary school pupils, cheerfulness and explaining the work was emphasized. (Morrison & McIntyre, 1969: p. 109)

These studies lead one to expect that students might be quite capable of choosing the teacher who most appropriately satisfied their expectations through his teaching behavior, and indeed trials of this in other areas, which generally have been unsystematic and not reported in the research literature, have found that pupils did indeed derive some satisfaction from the choice of teachers, just as university students do. At the university level, studies of student ratings of teachers have found that the fact that the course is elective, in other words that students chose a particular teachers, has a significant impact on the ratings of the teacher by the students. (Gage, 1972 p. 171)

The notion of matching developed by Hunt and his colleagues might more accurately be described as constructive mismatch. They reject simple matching in terms of personality:

We...disagree with some prevalent views of education, especially at the college level, which emphasize placing the student in the environment that is most congruent with his existing personality structure. In our view such procedures simply promote arrestation and thereby defeat the process of growth and progression, which should be the major goal of education. (harvey, Hunt, & Schroeder, 1961: p. 340)

Rather, their proposals rest on a value assertion: "abstract, conceptual structure and its associated characteristics of creativity, flexibility, stress tolerance, and broad-spectrum coping power is a desirable adaptive

state. This value assertion stems from a concern with the person's capacity to adapt to a changing environment". (Hunt, 1971: p. 18) This conception emphasizes the need for a student-environment match which provides a challenge to stimulate growth. Ideally, this could be provided in terms of a classroom of which the entire student population is homogeneous in stage of conceptual development". (Hunt, 1971: p. 25)

An earlier version of this work involved an exploratory study, intended to

Obtain some indication of the educational relevance of the model by exposing classroom groups of the same conceptual stage to the same educational environment. Students at a given stage are expected to share certain common characteristics based on conceptual level. They should respond favorably to certain forms of teaching and unfavorably to others, even though the environments may not necessarily be optimal for progression...in this study we ask, "do these classroom stage groups differ from one another in expected ways?" and if so, "do these differences make any educational sense to the teachers?" (1971: p. 26)

The success of the study lead to further theoretical work, and to the development of the most recent formulation, the Conceptual Level Matching Model, which involves specifying the desired change (educational objectives), a conception of the person (learner characteristics), a conception of the environment (educational approaches), and a conception of the interactive process (theory of instruction).

The early results of empirical work using this model have been quite interesting, and do suggest that the Conceptual Level of students can be a useful guideline for varying classroom groupings and educational approaches. Given that some teachers prefer different approaches to others, it also suggests the possibility of matching student requirements, based on conceptual

levels, and teacher predilections, based on preferred educational approaches. However, the tentativeness of the empirical results to date, and the relatively small samples involved, require that some caution be exercised in applying this particular notion of matching. It remains one of the more interesting new approaches, and should certainly generate a good many empirical studies.

Implementing the Proposal

A number of general points about evaluation and its impact on an educational system can be made, in the form of cautions to administrators wishing to implement an evaluation proposal such as this. The main point here is that evaluation in the context of educational organizations invariably has political implications, for the following reasons (based on Tumin, 1970):

1. It means different things to different people, because they have both different bases of judgment and aspire to different standards of competence.
2. Consensus agreements become a necessary way of proceeding with implementation and this consensus necessarily implies the need for some educational growth on the part of participants, and suggests that the whole evaluation process in an organization needs to develop by evolution rather than by fiat.
3. The evaluators need to be open to influence by all parties involved not forgetting students, in the planning and implementation stages particularly.

Evaluation is inevitably an exercise of power, and at the very least the power to discriminate levels of competence, and at the most the power to retain or release employees. It thus intrinsically threatening to all concerned, although there is some evidence that better than average and

non-tenured teachers have somewhat more positive attitudes towards evaluation than others. (Wagoner & O'Hanlon, n.d) Further, for students, evaluation may represent a shift in power in their favor: their participation in rating teachers represents a gain in net power. There is some evidence that this in itself is desirable, that

by giving some power to students, perhaps in the form of choice of teacher, or voluntary attendance, or monthly ratings of teacher performance one might well inculcate a higher priority for the goals of students and of the school system in the minds of teachers (Coleman, 1972: p. 46)

If these cautions are kept in mind, it is possible that the evaluation scheme can in fact contribute positively towards the general health of the school division as an organization. There is some evidence that the absence of control is relatively undesirable for all concerned. The desirable situation is participation in control. (See Tannenbaum, 1962) A study of teachers and bureaucracy found, contrary to expectation, that the sense of power was higher in highly bureaucratic schools: "Teachers in highly bureaucratic systems had a significantly higher, not lower, sense of power than those in less bureaucratic systems". (Moeller & Charters, 1966: p. 458)

With regard to more immediate problems of implementation, two related points can be made: the selection, assignment, and retention-release decisions regarding teachers will generally be made by a superintendent or assistant superintendent, and consequently implementation of the program proposed here will depend a good deal on the ability of these people to administer pencil-and-paper tests, and to carry out careful classroom observation. Fortunately, interaction analysis has been quite

highly developed, and instructional materials exist which make the skills readily teachable in in-service workshops. (see, for example, Amidon & Hunter, 1966)

The second point concerns in-service programs for teachers. At present these are rather poorly developed, primarily because there is no reasonable theoretical basis for these programs, and consequently they tend to be rather haphazard and only marginally effective. The technical skills approach to teacher training developed at Stanford University (Berliner, 1969) seems on the surface at least ideally suited for in-service work. The references here to evaluations used developmentally provide only very limited hints as to what the needs of teachers might be, but it does seem unlikely that a serious in-service program can be developed until after some relatively extensive and objective evaluation schemes have revealed the areas of greatest need. Teachers are not very likely to be aware of their deficiencies, although there is little doubt that student rating schemes could help to enlighten them. Thus the implementation of the teacher evaluation program suggested here is seen as a necessary prerequisite to the development of an adequate in-service program.

Implementing the proposal for improving teacher assignment requires rather more caution, in the opinion of the writer, than the evaluation-based proposals. Consulting teachers on their preferred assignment and consulting students on their preferred teacher both seem reasonable practices, eminently worth trying. However, reducing the discrepancies between actual and preferred educational contexts for teachers and students will not necessarily result in improved achievement. Careful evaluation of the outcomes would therefore be necessary. The Conceptual Level Model of Hunt and colleagues, which emphasizes constructive mismatch, seems likely

to be more productive of desirable student growth. However, this model should only be applied under carefully controlled conditions and preferably with the assistance of competent consultants. In essence, all of these matching possibilities require rather careful testing before they can be unequivocally recommended.

Conclusions and Implications

The approach to improving aggregate effectiveness proposed here purports to be timely in that a probable oversupply of teachers (for Manitoba, (see Husby, 1972; for Saskatchewan, see Scharf, 1972) makes it possible to replace release teachers, and also makes the problem of selecting and assigning teachers very much less difficult than previously. Rather than having one candidate who seems suitable for the required position, the administrator will probably have several candidates, and thus the problem of selection and assignment becomes one of applying more refined processes beyond the kind of gross matching which was possible in the past.

It is also maintained that the approach here is likely to be effective, in that the teacher behaviors and characteristics considered desirable here have had in empirical studies in the past positive relationships with good attitudes and achievement levels of students; unless evidence to the contrary exists, it seems reasonable to expect the positive outcomes to be replicated in Manitoba.

In addition, it is also maintained that the approach proposed here is practical, in that it can be utilized without extensive retraining of supervisors, or heavy expenditures for consultants and data analysis.

Unquestionably, some additional prospects for the in-service training of administrators are implied by the proposals made here; for example, few administrators in Manitoba have had exposure to verbal interaction analysis techniques for evaluating teachers, but this could easily be remedied in a school division wishing to adopt these proposals.

The proposal can conveniently be summarized in a simple table:

TEACHER EVALUATION AND ADMINISTRATIVE DECISION-MAKING:

DATA NEEDS AND ADMINISTRATOR RESPONSIBILITIES

<u>Administrative Decisions & Evaluation Data Needed</u>	<u>Sources of Data</u>	<u>Administrators Responsible for Collecting and Utilizing Data</u>
Selection (Predictive Data)	M.T.A.I. Verabal Facility Test	Superintendent Superintendent
Development (Formative Data)	Student Ratings Interaction Analysis Classroom Behavior	Teacher Principal Principal
Retention/Release (Summative Data)	Interaction Analysis Classroom Behavior M.T.A.I.	Principal, Superintendent Principal, Superintendent Superintendent

Additionally, it is proposed that matching models of three types are relevant to the improvement of teacher effectiveness: one, improving the match between teachers' preferences and training and their teaching assignments; two, improving the match between teacher and student expectations of role behaviors, and thus reducing role conflict; and three, improving the match between student needs for growth and teacher practices. The first two can be implemented to some extent by extending consultation and choice for

teachers and students.

Close attention by administrators to the processes of teacher selection, assignment, development, and retention/release can, it is maintained, substantially increase the aggregate teaching effectiveness of the staff of a school division over a period of some years. Serious administrative attempts to improve teaching have not been common in recent years. They may well become more popular in response to demands for accountability and to the improving supply of teachers.

LIST OF REFERENCES

- Alkin, M.C.
- 1972 "Accountability Defined" In, *UCLA Evaluation Comment*
Los Angeles: Center for the Study of Evaluation,
University of California.
- American Association of School Administrators
- 1972 "Evaluating Teaching Performance" In, *ERS Circular*,
Washington D.C.: Educational Research Service,
American Association of School Administrators.
- Amidon, E. & Hunter E.
- 1966 *Improving Teaching: The Analysis of Classroom Verbal
Interaction.* New York: Holt, Rinehart and Winston, Inc.
- Berliner, D.C.
- 1969 *Microteaching and the Technical Skills Approach to
Teacher Training.* Palo Alto: Stanford Center for
Research and Development in Teaching.
- Biddle, B.J.
- 1969 "Teacher Roles" In, *Encyclopedia of Educational Research*,
Fourth Edition, R.L. Ebel, (Ed.) New York: McMillan.
- Bogen, I.
- 1954 "Pupil-Teacher Rapport and the Teaching Awareness of
Status Structures Within the Group" In, *Journal of
Educational Sociology*, No. 28.
- Cheong, G.S. & DeVault, M.V.
- 1966 "Pupil Perceptions of Teachers" In, *Journal of Educational
Research*, No. 59.
- Coleman, P.
- 1970 "Organizational Effectiveness in Education: Its Measurement
and Enhancement" In, *Interchange* Vol. 3, No. 1.
- Corwin, R.G.
- 1970 *Militant Professionalism: A Study of Organizational
Conflict in High Schools.* New York: Appleton-Century-Crofts.

- Etzioni, A.
- 1972 "Human Beings Are Not Very Easy to Change After All." In, *The Saturday Review*, June 3rd.
- Flanders, N.A. & Simon, A.
- 1969 "Teacher Effectiveness" In, *Encyclopedia of Educational Research*, Fourth Edition, R.L. Ebel, (Ed.) New York: MacMillan.
- Gage, N.L.
- 1972 *Teacher Effectiveness and Teacher Education: The Search for a Scientific Basis.* Palo Alto: Pacific Books.
- Guthrie, J.W.
- 1970 "A Survey of School Effectiveness Studies". In, *Do Teachers Make A Difference?* Washington, D.C.: Bureau of Educational Personnel Development, Office of Education.
- Hanushek, E.
- 1970 The Production of Education, Teacher Quality, and Efficiency". In, *Do Teachers Make A Difference?* Washington, D.C.: Bureau of Educational Personnel Development, Office of Education.
- Harvey, O.J., Hunt, D.E. & Schroeder, H.M.
- 1961 *Conceptual Systems and Personality Organization.* New York: Wiley.
- Hunt, D.E.
- 1971 *Matching Models in Education.* Toronto: The Ontario Institute for Studies in Education.
- Husby, P.J.
- 1972 "Estimates of Student Enrolments and the Teaching Force in Manitoba Public Schools to 1979". In, *The Manitoba Journal of Education*, Vol. VIII, No. 2, Winnipeg: Manitoba Educational Research Council.
- Jackson, P.W.
- 1968 *Life in the Classroom.* New York: Holt, Rinehart and Winston, Inc.

- Kahn, R.L., Wolfe, D.M., Quinn, R.P. Snock, J.D., and Rosenthal, R.A.
 1964 *Organizational Stress: Studies in Role Conflict and Ambiguity.* New York: Wiley.
- Klein, S.P.
 1971 "The Uses and Limitations of Standardized Tests in Meeting the Demands for Accountability." In, *UCLA Evaluation Comment.* Los Angeles: Center for the Study of Evaluation, University of California.
- Levin, H.M.
 1970 "A Cost-Effectiveness Analysis of Teacher Selection". In, *The Journal of Human Resources*, Vol. 1.
- Moeller, G.H., & Charters, W.W.
 1966 "Relation of Bureaucratization to Sense of Power Among Teachers". In, *The Administrative Science Quarterly*, Vol. 10.
- Mood, A.M.
 1970 "Do Teachers Make A Difference?" In, *Do Teachers Make A Difference?* Washington, D.C.: Bureau of Educational Personnel Development, Office of Education.
- Morrison, A. & McIntyre, D.
 1969 *Teachers and Teaching.* Baltimore: Penguin Books.
- Popham, W.J.
 1968 *Performance Tests of Instructor Competence for Trade and Technical Education. Final Report.* Los Angeles University of California. ED 027 418.
- Rosenshine, B.
 1970 "Enthusiastic Teaching: A Research Review". In, *School Review.* No. 78.
- Ryans, D.G.
 1960 *Characteristics of Teachers.* Washington, D.C.: American Council on Education.

Scharf, M.P.

- 1972 "Implications of Population Trends for Rural Education." In, *The School Trustee*. Vol. XXV, No. 3, Regina: The Saskatchewan School Trustees Association.

Tannebaum, A.S.

- 1962 "Control in Organizations: Individual Adjustment and Organizational Performance". In, *The Administrative Science Quarterly*, Vol. 7.

Tumin, M.M.

- 1970 "Evaluation of the Effectiveness of Education: Some Problems and Prospects." In, *Interchange*, Vol. 1, No. 3.

Turner, R.L.

- 1963 "Task Performance and Teaching Skill in the Intermediate Grades". In, Silberman, H.F. (Ed.), *A Symposium on Current Research on Classroom Behavior of Teachers and Its Implications for Teacher Education*, *Journal of Teachers Education*. Vol. 14.

Wagoner, R.L. & O'Hanlon, J.P.

- n.d. *Teacher Attitude Toward Evaluation*. ED 013 236.

Wallin, H.A.

- 1971 *The Educational Employment Histories of the Professional Workforce in B.C. Schools, 1968-69*. Vancouver: Centre for the Study of Administration in Education, The University of British Columbia.

Yee, A.H.

- 1968 "Is the Minnesota Teacher Attitude Inventory Valid and Homogeneous?" In, *Journal of Educational Measurement*, Vol. 4.